Security analysis of an untrusted source for quantum key distribution: passive approach

2010 New J. Phys. 12 023024

(http://iopscience.iop.org/1367-2630/12/2/023024)

New Journal of Physics
The open-access journal for physics

Security analysis of an untrusted source for quantum key distribution: passive approach

Yi Zhao[1,2], **Bing Qi, Hoi-Kwong Lo and Li Qian**
Center for Quantum Information and Quantum Control, Department of Physics and Department of Electrical & Computer Engineering, University of Toronto, Toronto, Ontario M5S 3G4, Canada
E-mail: phy.zhao@utoronto.ca

New Journal of Physics **12** (2010) 023024 (33pp)
Received 30 May 2009
Published 16 February 2010
Online at http://www.njp.org/
doi:10.1088/1367-2630/12/2/023024

Abstract. We present a passive approach to the security analysis of quantum key distribution (QKD) with an untrusted source. A complete proof of its unconditional security is also presented. This scheme has significant advantages in real-life implementations as it does not require fast optical switching or a quantum random number generator. The essential idea is to use a beam splitter to split each input pulse. We show that we can characterize the source using a cross-estimate technique without active routing of each pulse. We have derived analytical expressions for the passive estimation scheme. Moreover, using simulations, we have considered four real-life imperfections: additional loss introduced by the 'plug&play' structure, inefficiency of the intensity monitor noise of the intensity monitor, and statistical fluctuation introduced by finite data size. Our simulation results show that the passive estimate of an untrusted source remains useful in practice, despite these four imperfections. Also, we have performed preliminary experiments, confirming the utility of our proposal in real-life applications. Our proposal makes it possible to implement the 'plug&play' QKD with the security guaranteed, while keeping the implementation practical.

[1] Author to whom any correspondence should be addressed.
[2] Present address: California Institute of Technology, MC 12-33, Pasadena, CA 91106, USA.

Contents

1. Introduction

Quantum key distribution (QKD) provides a means of sharing a secret key between two parties, a sender Alice and a receiver Bob, securely in the presence of an eavesdropper, Eve [1]–[3]. The unconditional security of QKD has been rigorously proved [4], even when implemented with imperfect real-life devices [5, 6]. The decoy state method was proposed [7]–[12] and experimentally demonstrated [13, 14] as a means to dramatically improve the performance of QKD with imperfect real-life devices with unconditional security still guaranteed [5, 9].

A large class of QKD setups adopts the so-called 'plug&play' architecture [15, 16]. In this setup, Bob sends strong pulses to Alice, who encodes her quantum information on them and attenuates these pulses to quantum level before sending them back to Bob. Both phase and polarization drifts are intrinsically compensated for, resulting in a very stable and relatively low quantum bit error rate (QBER). These significant practical advantages make the 'plug&play' very attractive. Indeed, most of the current commercial QKD systems are based on this particular scheme [17, 18].

The security of 'plug&play' QKD was a long-standing open question. A major concern arises from the following fact: when Bob sends strong classical pulses to Alice, Eve can freely manipulate these pulses, or even replace them with her own sophisticatedly prepared pulses. That is, the source is equivalently controlled by Eve in the 'plug&play' architecture.

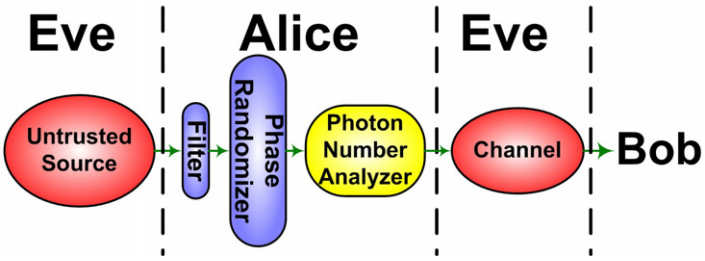

Figure 1. A general schematic of secure QKD with an unknown and untrusted source. The filter guarantees the single-mode assumption. The phase randomizer guarantees the phase randomization assumption. The photon number analyzer (PNA) estimates photon number distribution of the source. Various PNAs are shown in figure 2.

In particular, it is no longer correct to assume that the photon number distribution is Poissonian, as is commonly assumed in standard security proof. This is a major reason why standard security proofs such as GLLP [5] do not appear to apply directly to the 'plug&play' scheme.

It might be tempting to apply the central-limit theorem [19] to the current problem. That is, the photons contained in a pulse after heavy attenuation obey a Gaussian distribution asymptotically. The central-limit theorem was adopted in [20].

However, the central-limit theorem does not apply to the situation that the current paper is addressing. The current paper, as well as a previous work [21], does not rely on the central-limit theorem and removes the assumption on the input photon number distribution. That is, our analysis applies to sources with an arbitrary photon number distribution. For example, imagine a source that follows a dual-delta distribution (i.e. the pulses sent by the source contain either n_1 or n_2 photons, where n_1 and n_2 are large and different integers). In this case, even if Alice applies heavy attenuation on the input pulses, the resulting photon number per pulse distribution would be the sum of two Gaussian distributions, which in general is not a Gaussian distribution.

The dual-delta distributed source is of significant practical meaning rather than a purely imaginary source. Consider the case of the Trojan horse attack [20]: an eavesdropper occasionally sends a bright pulse to Alice and splits the corresponding output signal from Alice. In this case, the input photon number per pulse distribution on Alice's side would have two peaks: one corresponds to the photon number of the authentic source, and the other corresponds to the sum of the photon numbers from the two pulses (one from the authentic source and the other one from the eavesdropper's probing pulse). The security analysis that is based on the central-limit theorem (e.g. [20]) may not be directly applicable to this case. However, the analysis proposed in [21] and in the current paper can analyze such a case simply by defining an appropriate input photon number range of the untagged bits such that most input pulses are included. Note that the input photon number range for the untagged bits is defined in the post-processing stage, during which Alice has already collected the photon number distribution of the samples.

The unconditional security of the 'plug&play' QKD scheme has been recently proven in [21]. The basic idea is illustrated in figure 1. A filter guarantees the single-mode assumption. A phase randomizer guarantees the phase randomization assumption. Note that for the state that is accessible to the eavesdropper, Alice's phase randomization is equivalent to a quantum

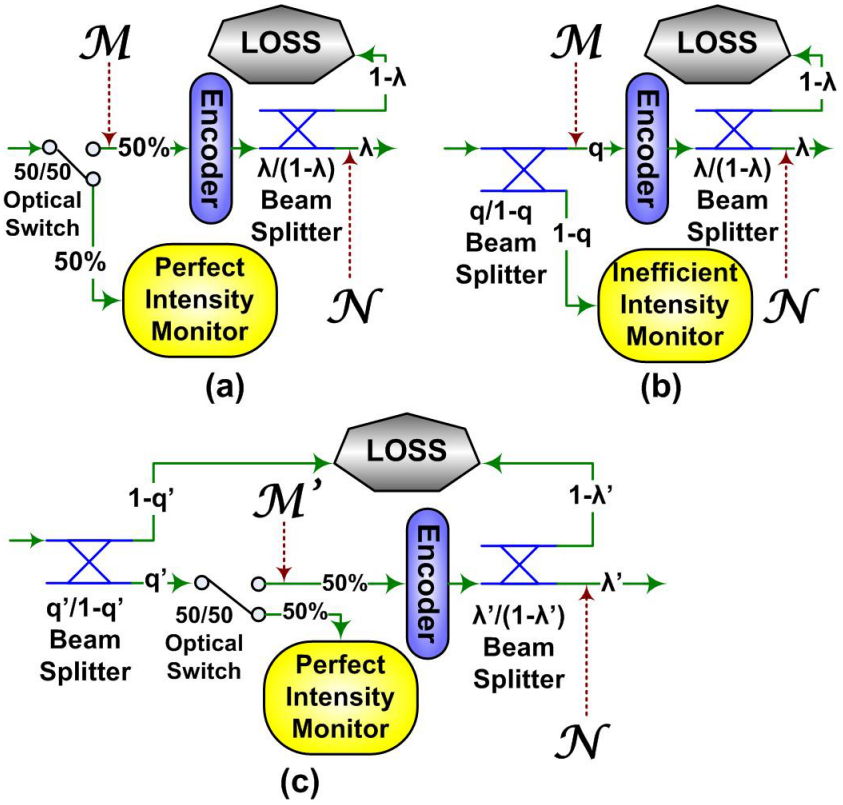

Figure 2. Different schemes to estimate photon number distribution. $\mathcal{M}$, $\mathcal{M}'$ and $\mathcal{N}$ are random variables for input photon number, virtual input photon number and output photon number, respectively. All the internal loss of Alice is modeled as a $\lambda/(1-\lambda)$ beam splitter (in (a) and (b)) or a $\lambda'/(1-\lambda')$ beam splitter (in (c)). (a) Active scheme, (b) passive scheme and (c) hybrid scheme. $q' = \eta_{\mathrm{IM}}(1-q)$, where $\eta_{\mathrm{IM}} \leqslant 1$ is the efficiency of the imperfect intensity monitor. $\lambda' = q\lambda/q'$. Note that the scheme shown in (c) is a virtual setup that has features from both the active scheme (a) and the passive scheme (b). The purpose of introducing this virtual scheme (c) is to bridge the active scheme (a) and the passive scheme (b).

non-demolition (QND) measurement of the photon numbers of the optical pulses. See appendix A for details. Therefore, from now on, without loss of generality, we will assume that Alice's input signal is a classical mixture of Fock states and, similarly, Alice's output signal is also a classical mixture of Fock states. A PNA estimates photon number distribution of the source. Details of the PNA in [21] are shown in figure 2(a).

The analysis presented in [21] applies to a general class of QKD with unknown and untrusted sources besides 'plug&play' QKD. For example, many QKD implementations use pulsed laser diodes as the light source. These laser diodes are turned on and off frequently to generate a laser pulse sequence. However, such laser pulses are not in a coherent state and the photon number per pulse does not obey Poisson distribution [21]. Moreover, the go-and-return scheme is also adopted by the recently proposed ground-satellite QKD project [22], in which the source is also equivalently unknown and untrusted.

Zhao *et al* [21] analyzes the photon number distribution of an untrusted source in the following manner: each input pulse will be randomly routed to either an encoder in figure 2(a) as a coding pulse or a perfect intensity monitor in figure 2(a) as a sampling pulse. The photon numbers of each sampling pulse are individually measured by the intensity monitor. In particular, one can obtain an estimate of the fraction of coding pulses that has a photon number $m \in [(1-\delta)M, (1+\delta)M]$ (here δ is a small positive real number and M is a large positive integer; both δ and M are chosen by Alice and Bob). These bits are defined as 'untagged bits'. The details of security analysis results of [21] are presented in appendix B. We note that some security analyses about QKD with a fluctuating source have been reported recently [23]–[26].

It is challenging and inefficient to implement the scheme proposed in [21], which is referred to as an active scheme, for the following reasons: (1) The optical switch in figure 2(a) is an active component and requires real-time control. The design and manufacture of the optical switch and its controlling system can be very challenging in high-speed QKD systems, which can operate as fast as 10 GHz [27]. (2) The random routing of optical pulses requires a high-speed sampling quantum random number generator (sampling QRNG), which does not yet exist for Gb/s systems. (3) The number of pulses sent to Bob is only a constant fraction (say half) of the number of pulses generated by the source, which means the key generation rate per pulse sent by the source is reduced by that fraction.

Naturally, the optical switch can be replaced by a beam splitter, which will passively split every input pulse, sending a portion into the intensity monitor and the rest to the encoder. This is referred to as a passive scheme. In this scheme, the sampling QRNG is not required.

A very recent work proposed some preliminary analysis on the passive estimation of an untrusted source using inverse Bernoulli transformation, and performed some experimental tests [28]. It is very encouraging to see that it is possible to prove the security of the passive estimate scheme for QKD with an untrusted source. As acknowledged by the authors of [28], the inverse Bernoulli transformation is beyond the computational power of current computers, and the required photon number resolution is beyond the capabilities of practical photodiodes. Owing to the above challenges, the experimental data reported in [28] were not analyzed by the analysis proposed in the same paper.

In this paper, we propose a passive scheme to estimate the photon number distribution of an untrusted source together with a complete proof of its unconditional security. We show that the unconditional security can still be guaranteed without routing each input optical pulse individually. Our analysis provides both an analytical method to calculate the final key rate and an explicit expression of the confidence level. Moreover, we considered the inefficiency and finite resolution of the intensity monitor, making our proposal immediately applicable. In the numerical simulation, we considered the additional loss introduced by the 'plug&play' structure and the statistical fluctuation introduced by the finite data size. We also gave examples of imperfect intensity monitors in the simulation, in which a constant Gaussian noise is considered.

This paper is organized as follows: in section 2, we propose a modified active estimate method; in section 3, we establish the equivalence between the modified active scheme proposed in section 2 and the passive estimate scheme; in section 4, we present a more efficient passive estimate protocol than the one proposed in section 3; in section 5, we present the numerical simulation results of the protocol proposed in section 4 and compare the efficiencies of active and passive estimates; in section 6, we present a preliminary experiment based on our proposed passive estimate protocol.

2. Modified active estimate

In [21], it is shown that Alice can randomly pick a fixed number of input pulses as sampling pulses, and measure the number of untagged sampling bits. One can then estimate the number of untagged coding bits.

We find that we can modify the scheme proposed in [21] by drawing a non-fixed number of input pulses as samples. A passive estimate can be built on top of this modified active estimate scheme. Note that we only modified the way to estimate the number of untagged coding bits. Once the number of untagged coding bits is estimated, the security analysis proposed in [21] is still applicable to calculate the lower bound of the secure key rate.

Lemma 1. *Consider that k pulses are sent to Alice from an unknown and untrusted source, within which V pulses are untagged. Alice randomly assigns each bit as either a sampling bit or a coding bit with equal probabilities (both are 1/2). In total, V_s sampling bits and V_c coding bits are untagged. The probability that $V_c \leqslant V_s - \epsilon k$ satisfies*

$$P(V_c \leqslant V_s - \epsilon k) \leqslant \exp\left(-\frac{k\epsilon^2}{2}\right), \tag{1}$$

where ϵ is a small positive real number chosen by Alice and Bob.

That is, Alice can conclude that $V_c > V_s - \epsilon k$ with confidence level

$$\tau > 1 - \exp\left(-\frac{k\epsilon^2}{2}\right). \tag{2}$$

Proof. See appendix C. □

Note that the right-hand side of equation (1) is independent of V. This is important because Alice does not know the exact value of V, whereas Eve may know and may even manipulate the value of V. Nonetheless, the inequality suggested in equation (1) holds for any possible value of V. Therefore, Alice can always estimate that $V_c > V_s - \epsilon k$ with confidence level $\tau_a \geqslant 1 - \exp(-k\epsilon^2/2)$. Note that the estimate given in lemma 1 is actually quite good for us because we will mainly be interested in the case where V is close to k.

3. From active estimate to passive estimate

The PNA of our proposed scheme is shown in figure 2(b) and the entire scheme is shown in figure 3. We replaced the 50/50 optical switch in figure 2(a) by a $q/(1-q)$ beam splitter in figure 2(b). In this scheme, each input pulse is passively split into two: one (defined as U pulse) is sent to the encoder and transmitted to Bob and the other (defined as L pulse) is sent to the intensity monitor. The visualization of U/L pulses is shown in figure 4.

One may naïvely think that since the beam splitting ratio q is known, one can easily estimate the photon number of the U pulse from the measurement result of photon number of the corresponding L pulse. However, this is not true. Any input pulse, after the phase randomization, is in a number state. Therefore, for a pair of U and L pulses originating from the same input pulse, the total photon number of the two pulses is an unknown constant. This restriction suggests that we should not treat the photon numbers of two such pulses as independent variables, and the random sampling theorem cannot be directly applied.

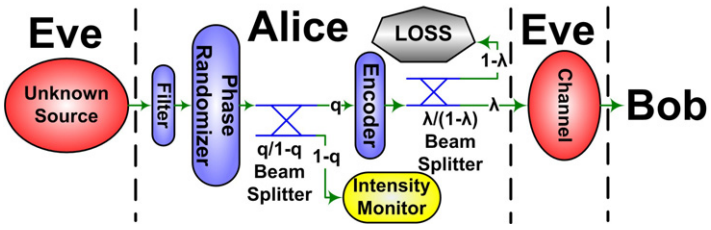

Figure 3. A schematic diagram of our proposed secure QKD scheme with a passive estimate on an unknown and untrusted source. The filter guarantees the single-mode assumption, and the $q/(1-q)$ beam splitter and the intensity monitor are used to passively estimate the photon number of input pulses. All the internal losses inside Alice's local lab are modeled as a $\lambda/(1-\lambda)$ beam splitter. That is, any input photon has λ probability to get encoded and sent from Alice to Bob and $1-\lambda$ probability to be lost.

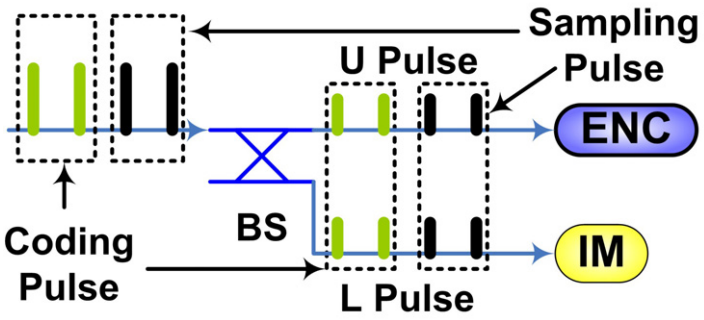

Figure 4. Visualization of different types of pulses. BS: beam splitter; ENC: encoder; IM: intensity monitor. Each input pulse is randomly assigned as either a coding pulse or a sampling pulse. After entering the beam splitter, each pulse is split into a U pulse that enters the encoder and an L pulse that enters the intensity monitor. As a result, there are four types of pulses: coding U pulse, coding L pulse, sampling U pulse and sampling L pulse.

To bridge the active scheme (in figure 2(a)) and the passive scheme (in figure 2(b)), we introduce a virtual setup (in figure 2(c)). We call such a virtual setup a 'hybrid' scheme because it has features from both the active and the passive schemes.

We assume that the inefficiency of the intensity monitor can be modeled as an additional loss [28]. In the passive scheme (figure 2(b)), assuming that the efficiency of the intensity monitor is $\eta_{\mathrm{IM}} \leqslant 1$, the probability that an input photon is detected is

$$q' = (1-q)\eta_{\mathrm{IM}}. \tag{3}$$

Therefore, we could model the $q/(1-q)$ beam splitter and the inefficient intensity monitor in figure 2(b) as a $q'/(1-q')$ beam splitter and a perfect intensity monitor as in figure 2(c).

The above modification changes the probability that an input photon is sent to Bob. To ensure that an identical attenuation is applied to the coding pulses in both the passive scheme (in figure 2(b)) and the hybrid scheme (in figure 2(c)), we re-define the internal transmittance in the virtual setup as

$$\lambda' = q\lambda/q' \leqslant 1. \tag{4}$$

For a given input photon number distribution, the output photon number distribution is determined by the internal loss [21]. Since the internal losses in the passive scheme and the hybrid scheme are identical, for a given input photon number distribution (which can be unknown), the output photon number distribution of the passive scheme is identical to that of the hybrid scheme. Moreover, the photon number distributions obtained by the intensity monitors are also identical for these two schemes.

Note that this virtual setup is not actually used in an experiment, but is purely for building the equivalence between the active and the passive schemes.

By putting equations (3) and (4) together, we have one constraint:

$$\lambda' = \frac{q\lambda}{(1-q)\eta_{\mathrm{IM}}} \leqslant 1. \tag{5}$$

This constraint is very easy to meet in an actual experiment as λ can be lower than 10^{-6} in a practical setup [21], $q/(1-q) \leqslant 100$ in typical beam splitters and η_{IM} can be greater than 50% in commercial photodiodes[3].

The resolution of the intensity monitor is another important imperfection. In a real experiment, the intensity monitor may indicate that a certain pulse contains m' photons. Here we refer to m' as the *measured* photon number in contrast to the *actual* photon number m. However, due to the noise and the inaccuracy of the intensity monitor, this pulse may not contain exactly m' photons. To quantify this imperfection, we introduce a term 'the conservative interval' ς. We then define $\underline{V}^{\mathrm{L}}$ as the number of L pulses with measured photon number $m' \in [(1-\delta)M + \varsigma, (1+\delta)M - \varsigma]$. One can conclude that, with confidence level $\tau_{\mathrm{c}} = 1 - c(\varsigma)$, the number of untagged L bits $V^{\mathrm{L}} \geqslant \underline{V}^{\mathrm{L}}$. One can make $c(\varsigma)$ arbitrarily close to 0 by choosing large enough ς[4]. The conservative interval is a statistical property rather than an individual property. That is, for one individual pulse, the probability that $|m - m'| > \varsigma$ can be non-negligible.

In the virtual setup, input pulses are treated in the same manner as in the active estimate scheme: coding pulses are routed to the encoder and then sent to Bob, whereas the sampling pulses are routed to the perfect intensity monitor to measure their photon numbers. We can use the measurement results of sampling pulses to estimate the number of untagged bits in the coding pulses. Knowing the number of untagged bits, one can easily calculate the upper and lower bounds of the output photon number probabilities [21].

Since the passive scheme and the hybrid scheme share the same source, the output photon number distribution is solely determined by the internal loss. The internal transmittances for the coding bits are the same ($q'\lambda' = q\lambda$) for both schemes. Therefore, the upper and lower bounds of output photon number probabilities estimated from the hybrid scheme are also valid for those of the passive scheme.

Corollary 1. *Consider k pulses sent from an unknown and untrusted source to Alice, where k is a large positive integer. Alice randomly assigns each input pulse as either a sampling pulse or a coding pulse with equal probabilities. Define variables* $V_{\mathrm{s}}^{\mathrm{L}}$ *and* $V_{\mathrm{c}}^{\mathrm{U}}$ *as the number of untagged*

[3] Several commercial high-speed InGaAs photodiodes, including Thorlabs FGA04, JDSU EPM745 and Hamamatsu G6854-01, are claimed to have conversion efficiency over 70% at 1550 nm.

[4] The specific expression of $c(\varsigma)$ depends on the properties of a specific intensity monitor. Nonetheless, one can always make $c(\varsigma)$ arbitrarily close to 0 by choosing a large enough ς. That is, $\forall \zeta > 0$, we can always find $|\underline{\varsigma} \in [0, \delta M]|$ such that for any $|\varsigma \geqslant \underline{\varsigma}|$, we have $c(\varsigma) < \zeta$. Note that $c(\delta M) = 0$.

sampling L pulses and the number of untagged coding U pulses, respectively. Here U pulses are defined as pulses sent to the encoder in figure 4, and L pulses are defined as pulses sent to the intensity monitor in figure 4. Alice can conclude that $V_c^U > V_s^L - \epsilon_1 k$ *with confidence level* $\tau_1 \geqslant 1 - e^{-k\epsilon_1^2/2}$. *Here* ϵ_1 *is a small positive real number chosen by Alice and Bob. To calculate the upper and lower bounds of output photon number probabilities, one should use equivalent internal transmittance* λ', *which is given in equation (5), instead of actual internal transmittance* λ.

Proof. The sampling L pulses are sent to a perfect intensity monitor with a probability $q' = q\eta_{\mathrm{IM}}$. If we apply the same transmittance q' to the coding U pulses, we can consider sampling L pulses and the coding U pulses as a group of pulses that go through the same attenuation, and we randomly assign each pulse in the group as either a sampling L pulse or a coding U pulse with equal probabilities. Therefore, one can conclude that $V_c^U > V_s^L - \epsilon_1 k$ with confidence level $\tau_1 \geqslant 1 - e^{-k\epsilon_1^2/2}$ by applying Lemma 1. Since the overall transmittance for the U pulses is $q\lambda$, the internal transmittance for the untagged coding U pulses should be considered as $\lambda' = q\lambda/q'$. □

There is no physical location (e.g. between the beam splitter and the encoder in figure 4) where the U pulses see a transmittance of q' in the passive scheme. The output photon number probabilities of the coding U pulses are analyzed in the following manner: The coding U pulses, after propagating through a virtual transmittance q', contain V_c^U untagged bits. These coding U pulses then propagate through another virtual transmittance λ', and we can calculate the output photon number probabilities, which are identical to the output photon number probabilities generated by sending the coding U pulses through the real transmittance $q\lambda = q'\lambda'$.

Note that it is not clear to us how to use the random sampling theorem to estimate the number of untagged *coding* 'U' pulses from the number of untagged *coding* 'L' pulses. This is due to the correlations between corresponding 'L' and 'U' pulses. As discussed before, their photon numbers are not independent variables. We are applying a restricted sampling where we draw only one sample from each pair of U and L pulses.

A common imperfection is the inaccuracy of the beam splitting ratio q. One can calibrate the value of q, but only with a finite resolution. In the security analysis, one should pick the most conservative value of q within the calibrated range, that is, the value of q that suggests the lowest key generation rate. A similar strategy should be applied to the inaccuracy of internal transmittance λ.

4. Efficient passive estimate on an untrusted source

In the above analysis, only half pulses (coding pulses) are used to generate the secure key. Note that we can also use the measurement result of coding 'L' pulses to estimate the number of untagged sampling 'U' pulses as there is no physical difference between sampling pulses and coding pulses. Note that Alice has the knowledge of the number of untagged coding 'L' pulses. We have the following statement:

Corollary 2. *Consider k pulses sent from an unknown and untrusted source to Alice, where k is a large positive integer. Alice randomly assigns each input pulse as either a sampling pulse or a coding pulse with equal probabilities. Define variables* V_c^L *and* V_s^U *as the number of untagged coding L pulses and the number of untagged sampling U pulses, respectively. Here U pulses*

are defined as pulses sent to the encoder in figure 4, and L pulses are defined as pulses sent to the intensity monitor in figure 4. Alice can conclude that $V_s^U > V_c^L - \epsilon_2 k$ *with confidence level* $\tau_2 \geqslant 1 - e^{-k\epsilon_2^2/2}$. *Here* ϵ_2 *is a small positive real number chosen by Alice and Bob.*

A natural question is: since Alice has knowledge about both V_s^L and V_c^L, how can she estimate the number of total untagged U pulses, V^U $(= V_s^U + V_c^U)$?

Combining all untagged U bits is not entirely trivial. Consider that the untrusted source generates k pulses. Each of them is divided into two pulses. Therefore Alice and Bob have $2k$ pulses to analyze. However, these $2k$ pulses are *not* independent because the beam splitter clearly creates correlations between the corresponding L pulse and U pulse. A naïve application of the random sampling theorem, ignoring the correlation between U pulses and L pulses, may lead to a security loophole.

Lemma 2. *Consider k pulses sent from an unknown and untrusted source to Alice. Alice randomly assigns each input pulse as either a sampling pulse or a coding pulse with equal probabilities. Each input pulse is split into a U pulse and an L pulse (see figure 4 for visualization). The probability that* $V^U \leqslant V_s^L + V_c^L - \epsilon_1 k - \epsilon_2 k$ *satisfies*

$$P(V^U \leqslant V_s^L + V_c^L - (\epsilon_1 + \epsilon_2)k) \leqslant \exp\left(\frac{-k\epsilon_1^2}{2}\right) + \exp\left(\frac{-k\epsilon_2^2}{2}\right). \tag{6}$$

Proof. See appendix D. □

In real experiment, it is convenient to count *all* the untagged L pulses, defined as variable V^L $(= V_s^L + V_c^L)$. Can we estimate V^U directly from V^L?

Proposition 1. *Consider k pulses sent from an unknown and untrusted source to Alice. Alice randomly assigns each input pulse as either a sampling pulse or a coding pulse with equal probabilities. The probability that* $V^U \leqslant V^L - \epsilon k$ *satisfies*

$$P(V^U \leqslant V^L - \epsilon k) \leqslant 2\exp\left(\frac{-k\epsilon^2}{4}\right). \tag{7}$$

That is, Alice can conclude that $V^U > V^L - \epsilon k$ *with confidence level*

$$\tau > 1 - 2\exp\left(\frac{-k\epsilon^2}{4}\right). \tag{8}$$

Proof. This is a natural conclusion from Lemma 2. Note that $V^L = V_s^L + V_c^L$. If Alice chooses $\epsilon_1 = \epsilon_2 = \epsilon/2$, equation (6) reduces to equation (7). □

Once the number of untagged bits that are sent to Bob is estimated, the final key generation rate can be calculated [21].

5. Numerical simulation

We performed numerical simulation to test the efficiencies of the active and passive estimates. Here, we define the key generation rate as secure key bits per pulse sent by the *source*, which

Table 1. Simulation parameters from GYS [29].

η_{det}	α	Y_0	e_{det}
4.5%	0.21 dB km^{-1}	1.7×10^{-6}	3.3%

may be controlled by an eavesdropper. This is different from the definition used in [21], where the key generation rate is defined as secure key bits per pulse sent by *Alice*. Note that, in the passive scheme, *all* the pulses sent by the source are sent from Alice to Bob, whereas in the active scheme, only *half* of the pulses sent by the source are sent from Alice to Bob. Therefore, for the same setup, we can expect the key generation rate suggested by the passive scheme to be roughly twice as high as that by the active scheme. However, the equivalent input photon number in the passive scheme is lower than that of the active scheme, which introduces a competing factor. The comparison between passive and active estimates is discussed in the following sections.

5.1. Simulation techniques

The simulation technique in this paper is similar to that presented in [21] with a few improvements. Here we briefly reiterate it: firstly, we simulate the experimental outputs based on the parameters reported by Gobby *et al* [29], which are shown in table 1. At this stage, we assume that the source is Poissonian with an average output photon number M. For a QKD setup with channel transmittance η ($= \mathrm{e}^{-\alpha l}$, where α is the fiber loss coefficient and l is the fiber length between Alice and Bob), Bob's quantum detection efficiency η_{Bob}, detector intrinsic error rate e_{det} and background rate Y_0, the gain[5] and the QBER of the signals are expected to be [10]

$$\begin{aligned} Q_{\text{e}} &= Y_0 + 1 - \exp(-\eta\eta_{\text{Bob}} M\lambda), \\ E_{\text{e}} &= \frac{e_0 Y_0 + e_{\text{det}}[1 - \exp(-\eta\eta_{\text{Bob}} M\lambda)]}{Q_{\text{e}}}, \end{aligned} \tag{9}$$

respectively. Here Q_{e} and E_{e} refer to the experimentally measured overall properties rather than the properties of the untagged bits. Secondly, we calculate the secure key generation rate. The general expression of the secure key generation rate per pulse sent by Alice is given by [5, 9]

$$R \geqslant \tfrac{1}{2}[-Q_{\text{e}} f(E_{\text{e}}) H_2(E_{\text{e}}) + Q_1(1 - H_2(e_1))], \tag{10}$$

where f ($\geqslant 1$) is the bi-directional error correction inefficiency ($f = 1$ iff the error correction procedure achieves the Shannon limit), H_2 is the binary Shannon entropy, Q_1 is the gain of the single photon state in untagged bits and e_1 is the QBER of the single photon state in untagged bits. Q_{e} and E_{e} can be experimentally measured. Here, we use equations (9) to simulate the experimental outputs.

Q_1 and e_1 need to be estimated. Here, we use the method described in appendix B. The key assumption for decoy state QKD with an untrusted source is that $Y_{m,n}$ is identical for different states, and so is $e_{m,n}$ [21]. Here $Y_{m,n}$ is the conditional probability that Bob's detectors click

[5] The gain is defined to be the ratio of the number of receiver Bob's detection events to the number of signals emitted by sender Alice in the cases where Alice and Bob use the same basis. It depends mainly on the intensity of signal, channel transmittance and Bob's quantum efficiency.

given that this bit enters Alice's lab with photon number m and emits from Alice's lab with photon number n, and $e_{m,n}$ is the QBER of bits with m input photons and n output photons.

At the second stage, we do not make any assumption about the source. That is, Alice and Bob have to characterize the source from the experimental output. Note that we need to set the values of λ and δ (recall that all untagged bits have input photon numbers $m \in [(1-\delta)M, (1+\delta)M]$, where δ is a small positive real number, M is a large positive integer and both δ and M are chosen by Alice and Bob). It is preferable to set λ and δ to the values that yield the highest final key generation rate. We optimize the values of λ and δ numerically by exhaustive search. Moreover, in the simulation of decoy state QKD with a finite data size, we also need to optimize the portion of each state.

As a clarification, our security analysis does *not* require any additional assumptions of the source to analyze *experimental* outputs.

An important improvement is that the value of δ is optimized at all distances in the following simulations, while δ is set to be constant in [21]. This is because for different channel losses, the optimal value of δ can vary. Moreover, several important practical factors are considered, including the unique characteristic of plug&play structure, intensity monitor imperfections and finite data size.

For ease of calculation, similar to in [21], we approximate the Poisson distribution as a Gaussian distribution centered at M with variance $\sigma^2 = M$. This is an excellent approximation because M is very large (10^3 or larger) in all the simulations presented below.

There are various types of imperfections and errors. We will consider them one by one in the following sections. In section 5.3, we consider the asymmetry of the beam splitter. In section 5.4, we consider the source attenuation introduced by the bi-directional scheme. In section 5.5, we consider the inefficiency and the inaccuracy of the intensity monitor. In section 5.6, we consider the statistical fluctuation due to a finite data size.

5.2. *Infinite data size with perfect intensity monitor*

In the asymptotic case, Alice sends infinitely many bits to Bob (i.e. $k \to \infty$). Therefore we can set $\epsilon \to 0$ while still having $\tau \to 1$.

We assume that the intensity monitor is efficient and noiseless. Similarly to the case in [21], we set $M = 10^6$. Moreover, we set $q = 0.5$ as the 50/50 beam splitter is widely used in many applications.

The simulation results of the GLLP protocol [5], weak + vacuum decoy state protocol [10] and one-decoy protocol [10] are shown in figures 5, 6 and 7, respectively. We can see that the key generation rate of the passive estimate scheme on an untrusted source is very close to that on a trusted source, whereas the key generation rate of the active estimate scheme is roughly $1/2$ of that of the passive scheme. This is expected because in the active scheme, only half of the pulses generated by the source are sent to Bob, whereas in the passive scheme, all the pulses generated by the source are sent to Bob. Note that, in the asymptotic case, the efficiency of the active estimate scheme can be doubled by sending most pulses (asymptotically all the pulses) to Bob. In this case, there are still infinitely many pulses sent to the intensity monitor.

For ease of discussion, in the passive estimate scheme, we define untagged bits as bits with input photon number $m_{\mathrm{p}} \in [(1-\delta_{\mathrm{p}})M_{\mathrm{p}}, (1+\delta_{\mathrm{p}})M_{\mathrm{p}}]$, whereas in the active estimate scheme, we define untagged bits as bits with input photon number $m_{\mathrm{a}} \in [(1-\delta_{\mathrm{a}})M_{\mathrm{a}}, (1+\delta_{\mathrm{a}})M_{\mathrm{a}}]$. Here δ_{p} and δ_{a} are small positive real numbers chosen by Alice and Bob, and M_{p} and M_{a} are large

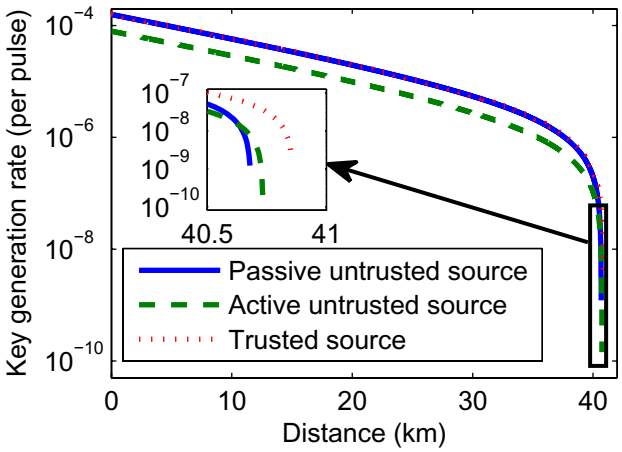

Figure 5. Simulation result of GLLP [5] protocol with infinite data size, symmetric beam splitter, perfect intensity monitor and uni-directional structure. We assume that the source is Poissonian centered at $M = 10^6$ photons per pulse, and the beam splitting ratio $q = 0.5$. See experimental parameters in table 1. We calculated the ratio of the key generation rate with an untrusted source over that with a trusted source. For the passive estimate scheme, the ratios are 98.4, 98.1 and 79.8% at 1, 20 and 40 km, respectively. For the active estimate scheme, the ratios are 49.4, 49.3 and 42.8% at 1, 20 and 40 km, respectively.

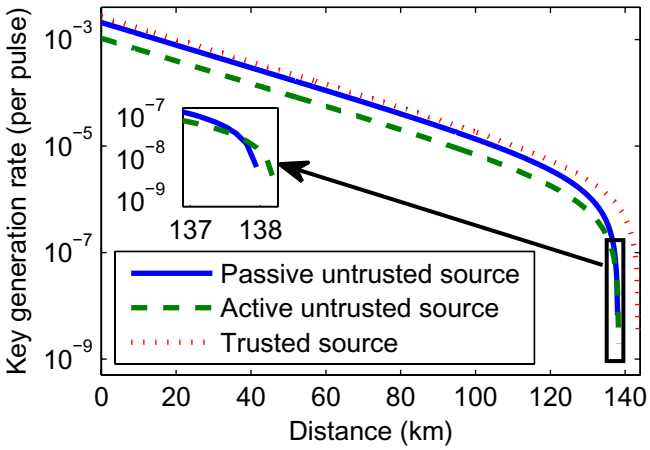

Figure 6. Simulation result of weak + vacuum [10] protocol with infinite data size, symmetric beam splitter, perfect intensity monitor and uni-directional structure. We assume that the source is Poissonian centered at $M = 10^6$ photons per pulse, and the beam splitting ratio $q = 0.5$. See experimental parameters in table 1. We calculated the ratio of the key generation rate with an untrusted source over that with a trusted source. For the passive estimate scheme, the ratios are 77.7, 77.1 and 73.8% at 1, 50 and 100 km, respectively. For the active estimate scheme, the ratios are 39.2, 39.0 and 37.4% at 1, 50 and 100 km, respectively.

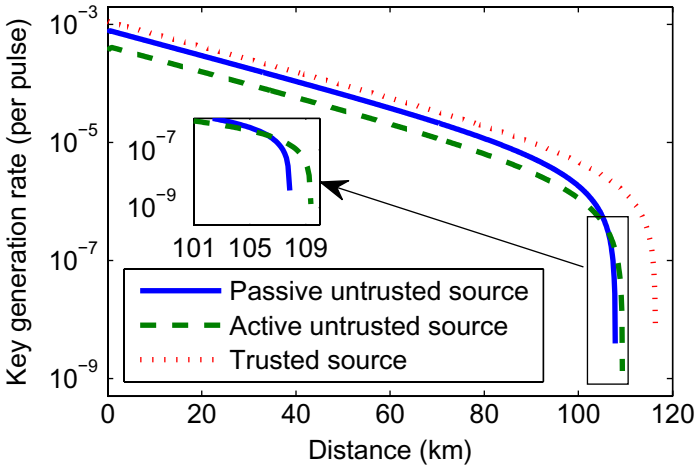

Figure 7. Simulation result of the one-decoy [10] protocol with infinite data size, symmetric beam splitter, perfect intensity monitor and uni-directional structure. We assume that the source is Poissonian centered at $M = 10^6$ photons per pulse, and the beam splitting ratio $q = 0.5$. See experimental parameters in table 1. We calculated the ratio of the key generation rate with an untrusted source over that with a trusted source. For the passive estimate scheme, the ratios are 71.5, 68.6 and 39.5% at 1, 50 and 100 km, respectively. For the active estimate scheme, the ratios are 38.0, 36.7 and 24.4% at 1, 50 and 100 km, respectively.

positive integers chosen by Alice and Bob. In the passive estimate scheme, we define the maximum possible tagged ratio as Δ_p. In the active estimate scheme, we define the maximum possible tagged ratio as Δ_a. Here the tagged ratio is defined as the ratio of the number of tagged bits over the number of all the bits sent to Bob.

By magnifying the tails at long distances (shown in the insets of figures 5–7), we can see that the active schemes suggest a higher key generation rate than the passive schemes do in all three protocols. This behavior is related to the following fact: in the passive estimate scheme, the equivalent input photon number is lower than that of the active estimate scheme. This is because the input photon number is defined as the photons counted by the intensity monitor, and only a portion of an input pulse is sent to the intensity monitor in the passive scheme. Compared to the active scheme, lower input photon number in the passive scheme leads to a larger coefficient of variation of measured input photon number distribution, assuming the source is Poissonian. Therefore, for the same source, if one sets $\delta_p = \delta_a$, Δ_p will be greater than Δ_a.[6] Increasing the coefficient of variation of the measured input photon number distribution will in general deteriorate the efficiency of the estimate for QKD with untrusted sources. Take two extreme cases for example: if the coefficient of variation is very large, which means the input photon number distribution is almost a uniform distribution, then the estimate efficiency will be very poor because either δ or Δ (or both) will be very large. If the coefficient of variation

[6] The values of δ in the passive estimate and the active estimate schemes are optimized separately in our simulation. The optimal value of δ_p usually deviates from the optimal value of δ_a with the same experimental parameters. Here we cite '$\delta_p = \delta_a$' just to illustrate an intuitive understanding of the phenomena shown in the insets of figures 5–7.

is very small, which means the input photon number distribution is almost a delta function, then the estimate efficiency will be very good because both δ and Δ can be very small.

The estimate of the gain of untagged bits is very sensitive to the value of Δ, especially when the experimentally measured overall gain is small (i.e. when the distance is long, which corresponds to the tails of figures 5–7). The estimate of untagged bits' gain is discussed in section 3 of [21]. Here we briefly recapitulate the main idea: Alice cannot, in practice, perform a QND measurement on the photon numbers of input pulses. Therefore, Alice and Bob do not know which bits are tagged and which are untagged, although they can estimate the minimum number of untagged bits. Without knowing which bits are untagged, Alice and Bob cannot measure the exact gain Q of untagged bits. Alice and Bob can only experimentally measure the overall gain Q_e, which contains contributions from both tagged bits and untagged bits.

Alice and Bob can still estimate the upper and lower bounds of Q. They can first estimate the maximum tagged ratio Δ. This estimate can be obtained either actively as proposed in [21] or passively as discussed in this paper. Alice and Bob can then estimate the upper and lower bounds of Q as follows [21]:

$$\begin{aligned} \overline{Q} &= \frac{Q_e}{1-\Delta-\epsilon}, \\ \underline{Q} &= \max\left(0, \frac{Q_e-\Delta-\epsilon}{1-\Delta-\epsilon}\right). \end{aligned} \tag{11}$$

$\underline{Q}$ is very sensitive to Δ when Q_e is small. Therefore, when the distance is long (which corresponds to the tails of figures 5–7), Q_e becomes very small, and $\underline{Q}$ will then be very sensitive to Δ. Since $\Delta_p > \Delta_a$, the passive estimate becomes less efficient than the active estimate in this case.

On the other hand, in short distances, Q_e is significantly greater than Δ_p and Δ_a; therefore the difference between Δ_p and Δ_a makes a negligible contribution to the performance difference between the passive and active estimates. At short distances, it is the following fact that dominates the performance difference between these two schemes: the passive estimate scheme can send Bob twice as many pulses as the active estimate scheme can.

One can increase δ to decrease Δ_p. That is, if one intends to ensure that $\Delta_p = \Delta_a$, one has to set $\delta_p > \delta_a$. However, increasing δ also has a negative effect on the key generation rate. This is discussed in sections 3 and 4 of [21].

In brief, lower input photon number is the reason why the passive estimate suggests a lower key generation rate than the active estimate does around maximum transmission distances in all of the three simulated protocols. This will be confirmed in the simulation presented in sections 5.3–5.6.

5.3. Biased beam splitter

A natural measure to improve the efficiency of the passive estimate is to increase input photon number. Note that in the passive estimate, as discussed in section 3, input photon numbers are the photon numbers counted by the intensity monitor. Therefore, it can improve the passive estimate's efficiency to send more photons to the intensity monitor (i.e. setting q smaller).

To test this postulate, we performed another simulation to compare the performance of the passive estimate with different values of q. Similar to the above subsection, we assume that the

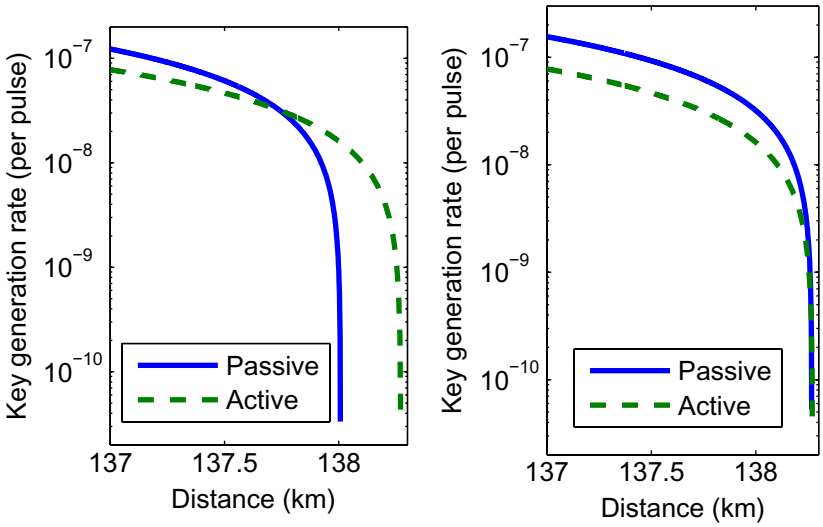

Figure 8. Simulation results for the weak + vacuum protocol [10] with different beam splitters for passive estimate. We assume that the data size is infinite, the intensity monitor is perfect, the source is Poissonian centered at $M = 10^6$ photons per pulse and the system is in uni-directional structure. See experimental parameters in table 1. The results are focused at the maximum transmission distance to illustrate the improvement of passive estimate by using a biased beam splitter that sends more photons into the intensity monitor. This is equivalent to increasing input photon numbers in the passive scheme.

intensity monitor is efficient and noiseless, and data size is infinite. Therefore $\epsilon = 0$. We set $M = 10^6$ at the *source*.

The simulation results are shown in figure 8. We can clearly see that by setting q to a smaller value (1%), the key generation rate of the passive estimate scheme is improved around the maximum transmission distance.

Intuitively, one can improve the efficiency of the active scheme by sending most pulses to Bob. One can refer to the discussion in appendix C below equation (C.4) as a starting point. Detailed discussion of optimizing the efficiency of the active estimate scheme is beyond the scope of the current paper and is subject to further investigation.

5.4. Plug&play setup

In the plug&play QKD scheme, the source is located in Bob's lab. Bright pulses sent by Bob will suffer the whole channel loss before entering Alice's lab. Therefore, in the plug&play setup, Alice's average input photon number is dependent of the channel loss between Alice and Bob. If the average photon number per pulse at the source in Bob's lab, M_B, is constant, the average input photon number per pulse in Alice's lab, M, decreases as the channel loss increases.

Similar to the above subsection, we assume that the intensity monitor is efficient and noiseless, and data size is infinite. Therefore $\epsilon = 0$. We set $M_B = 10^6$ at the source in Bob's lab. We set $q = 1\%$ to improve the passive estimate efficiency.

We clarify that 'distance' in all the simulations of the bi-directional QKD setup refers to a one-way distance between Alice and Bob, *not* a round-trip distance.

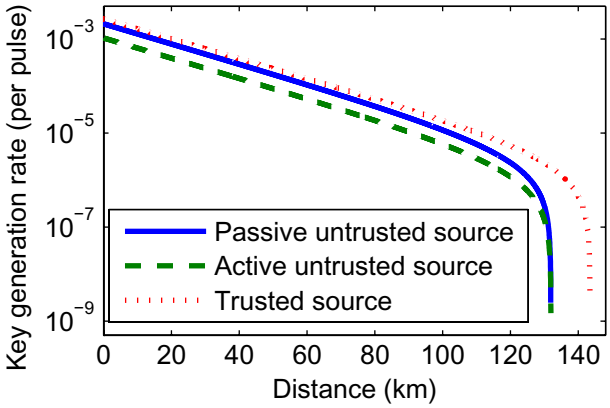

Figure 9. Simulation result of the weak + vacuum [10] protocol with infinite data size, asymmetric beam splitter, perfect intensity monitor and *bi-directional structure*. We assume that the source in Bob's lab is Poissonian centered at $M_B = 10^6$ photons per pulse, and the beam splitting ratio $q = 0.01$. See experimental parameters in table 1. We calculated the ratio of the key generation rate with an untrusted source over that with a trusted source. For the passive estimate scheme, the ratios are 78.5, 75.0 and 63.0% at 1, 50 and 100 km, respectively. For the active estimate scheme, the ratios are 39.2, 37.5 and 31.5% at 1, 50 and 100 km, respectively. Comparing with figure 6, we can see that the bi-directional nature of the plug&play setup reduced the efficiencies of both active and passive estimates on an untrusted source.

The simulation results of weak + vacuum protocol [10] are shown in figure 9. We can see that the bi-directional nature plug&play structure clearly deteriorates the performance at long distances for which the input photon number on Alice's side is largely reduced. This affects both the passive and active estimates.

A natural measure to improve the performance of the plug&play setup is to use a brighter source. By setting $M_B = 10^8$ at the source in Bob's lab, the performances for both passive and active estimates are improved substantially as shown in figure 10. Note that subnanosecond pulses with $\sim 10^8$ photons per pulse can be routinely generated with directly modulated laser diodes.

5.5. Imperfections of the intensity monitor

There are two major imperfections of the intensity monitor: inefficiency and noise. These imperfections are discussed in section 3. The inefficiency can be easily modeled as additional loss in the simulation.

There can be various noise sources, including thermal noise, shot noise, etc. Here, we consider a simple noise model where a *constant Gaussian* noise with variance σ_{IM}^2 is assumed. That is, if m photons enter an efficient but noisy intensity monitor, the probability that the measured photon number is m' obeys a Gaussian distribution

$$P_m(m') = \frac{1}{\sigma_{\mathrm{IM}}\sqrt{2\pi}} \exp\left[-\frac{(m-m')^2}{2\sigma_{\mathrm{IM}}^2}\right].$$

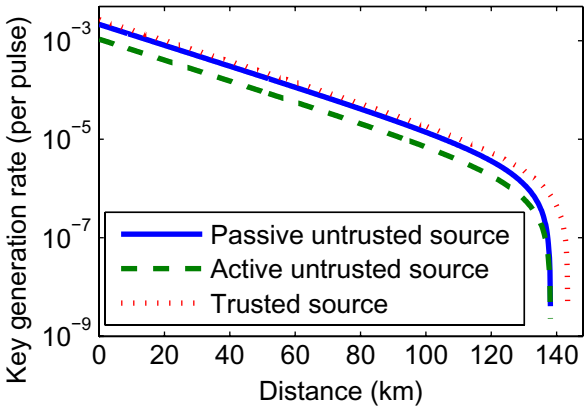

Figure 10. Simulation result of the weak + vacuum [10] protocol with infinite data size, asymmetric beam splitter, perfect intensity monitor, bi-directional structure and a bright light source. We assume that the source in Bob's lab is Poissonian centered at $M_B = 10^8$ photons per pulse, and the beam splitting ratio $q = 0.01$. See experimental parameters in table 1. We calculated the ratio of the key generation rate with an untrusted source over that with a trusted source. For the passive estimate scheme, the ratios are 80.3, 79.6 and 75.8% at 1, 50 and 100 km, respectively. For the active estimate scheme, the ratios are 40.1, 39.8 and 37.9% at 1, 50, and 100 km, respectively. Comparing with figure 9, we can see that the estimate efficiencies for both the active and passive schemes are improved by using a brighter source.

The measured photon number distribution $P(m')$ has larger variation than the actual photon number distribution $P(m)$ due to the noise of the intensity monitor. More concretely, if the actual photon numbers obey a Gaussian distribution centered at M with variance σ^2, the measured photon numbers also obey a Gaussian distribution centered at M, but with a variance $\sigma^2 + \sigma_{IM}^2$.

As in the previous subsections, we assume that the data size is infinite. Therefore $\epsilon = 0$. We set $M_B = 10^8$ at the source in Bob's lab. The plug&play setup is assumed. We set $q = 1\%$ to improve the passive estimate efficiency. The imperfections of the intensity monitor are set as follows: the efficiency is set as $\eta_{IM} = 0.7$, and the noise is set as $\sigma_{IM} = 10^5$ (see experimental parameters in sections 5.7 and 6). For ease of simulation, we assume that the intensity monitor conservative interval is constant[7] over different input photon numbers. We set $\varsigma = 6\sigma_{IM} = 6 \times 10^5$ to ensure a conservative estimate.

The simulation results for weak + vacuum protocol [10] are shown in figure 11. We can see that the detector noise significantly affects the performance of the plug&play QKD system. This is because at long distances, the bi-directional nature of the plug&play setup reduces the input photon number on Alice's side. Intensity monitor noise and the conservative interval are assumed as constants regardless of the input photon number in our simulation. Therefore they

[7] The assumption of constant conservative interval may not precisely describe the inaccuracy of the intensity monitor in realistic applications. Nonetheless, some factors, like finite resolution of analogue–digital conversion, may indeed be constant at different intensity levels. We remark that the noises of different intensity monitors may vary largely. Detailed investigation of the intensity monitor noise modeling is beyond the scope of the current paper.

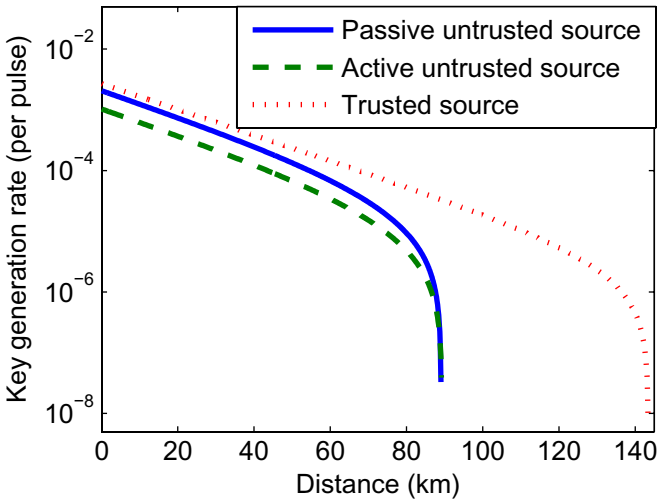

Figure 11. Simulation result of the weak + vacuum [10] *protocol with infinite data* size, asymmetric beam splitter, imperfect intensity monitor and bi-directional structure. We assume that the intensity monitor efficiency $\eta_{\mathrm{IM}} = 0.7$, the intensity monitor noise $\sigma_{\mathrm{IM}} = 10^5$, the intensity monitor conservative interval $\varsigma = 6 \times 10^5$, the source in Bob's lab is Poissonian centered at $M_{\mathrm{B}} = 10^8$ photons per pulse and the beam splitting ratio $q = 0.01$. See experimental parameters in table 1. Comparing with figure 9, we can see that the imperfections of the intensity monitor substantially reduce the efficiencies of both active and passive estimates.

become critical issues when the input photon number is low. As a result, the key generation rate at long distance is substantially reduced.

The above postulate is confirmed by the simulations shown in figures 12 and 13. In figure 12, we assume that the source in Bob's lab is extremely bright (sending out 10^{10} photons per pulse). We can see clearly that when the input photon number on Alice's side is high, the key generation rate is affected only slightly by the imperfections of the intensity monitor. Although it is challenging to build such bright pulsed laser diodes (10^{10} photons per pulse with pulse width less than 1 ns) at telecom wavelengths, one can simply attach a fiber amplifier to the laser diode to generate very bright pulses. Nonetheless, at such a high intensity level, nonlinear effects in the fiber, like self-phase modulation, may be significant [30].

An alternative solution is to use the uni-directional setting, in which the photon number per pulse is constantly high on Alice's side. From figure 13 we can see that using the uni-directional setting can also minimize the negative effects introduced by the imperfections of the intensity monitor. Nonetheless, if one adopts the uni-directional QKD scheme, one will lose the unique advantages of the bi-directional QKD scheme, like the intrinsic stability against the polarization dispersion and the phase drift. Note that adopting the uni-directional scheme does not mean that the coherent state assumption is valid. Indeed, even if Alice possesses the source, the source may not be Poissonian and Alice may not have a full characterization of the source without real-time monitoring.

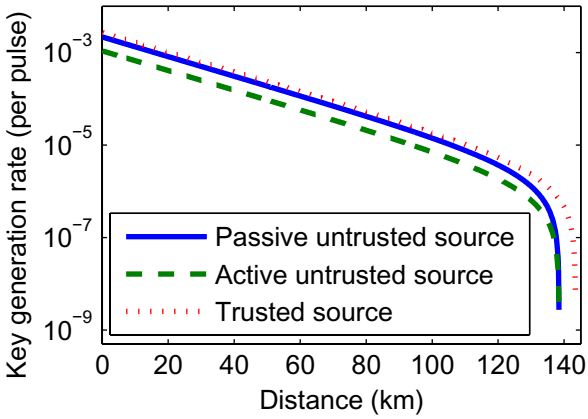

Figure 12. Simulation result of the weak + vacuum [10] protocol with infinite data size, asymmetric beam splitter, imperfect intensity monitor, bi-directional structure and a *very bright source*. We assume that the intensity monitor efficiency $\eta_{\mathrm{IM}} = 0.7$, the intensity monitor noise $\sigma_{\mathrm{IM}} = 10^5$, the intensity monitor conservative interval $\varsigma = 6 \times 10^5$, the source in Bob's lab is Poissonian centered at $M_{\mathrm{B}} = 10^{10}$ photons per pulse and the beam splitting ratio $q = 0.01$. Experimental parameters are cited in table 1. Comparing with figure 11, we can see that using a brighter source can effectively improve the efficiencies of both passive and active estimates. Although it is challenging to build such bright pulsed laser diodes (10^{10} photons per pulse with pulse width less than 1 ns) at telecom wavelengths, one can simply attach a fiber amplifier to the laser diode to generate very bright pulses. Nonetheless, at such a high intensity level, nonlinear effects in the fiber, like self-phase modulation, may be significant [30].

5.6. Finite data size

Real experiments are performed within a limited time, during which the source can only generate a finite number of pulses. To be consistent with previous analysis, we assume that the source generates k pulses in an experiment. Reducing the data size from infinite to finite has two consequences: Firstly, if the confidence level τ as defined in equation (8) (for passive estimate) or in equation (2) (for active estimate) is expected to be close to 1, ϵ has to be positive. More concretely, for a fixed k, if the estimate on the untrusted source is expected to have confidence level not less than τ, one has to pick ϵ as

$$\epsilon_{\mathrm{p}} = \sqrt{-\frac{4\ln((1-\tau)/(2))}{k}}$$

in the passive estimate scheme, or

$$\epsilon_{\mathrm{a}} = \sqrt{-\frac{2\ln(1-\tau)}{k}}$$

in the active estimate scheme. Secondly, in decoy state protocols [10], the statistical fluctuations of experimental outputs have to be considered. The technique to analyze the statistical fluctuation in decoy state protocols for numerical simulation is discussed in [10, 12, 14].

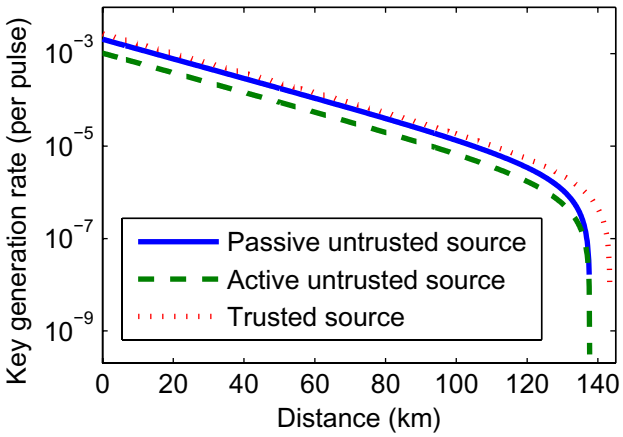

Figure 13. Simulation result of the weak + vacuum [10] protocol with infinite data size, asymmetric beam splitter, imperfect intensity monitor and *uni-directional structure*. We assume that the intensity monitor efficiency $\eta_{\mathrm{IM}} = 0.7$, the intensity monitor noise $\sigma_{\mathrm{IM}} = 10^5$, the intensity monitor conservative interval $\varsigma = 6 \times 10^5$, the source is Poissonian centered at $M = 10^8$ photons per pulse and the beam splitting ratio $q = 0.01$. Citing experimental parameters from table 1 and comparing with figure 11, we can see that uni-directional structure can effectively improve the efficiencies of both passive and active estimates.

In the simulation presented in figure 14, we assume that the data size is 10^{12} bits (i.e. the source generates 10^{12} pulses in one experiment). This data size is reasonable for the optical layer of the QKD system because reliable gigahertz QKD implementations have been reported in several recent works [27, 31, 32]. 10^{12} bits can be generated within a few minutes in these gigahertz QKD systems. We set the confidence level as $\tau \geqslant 1\text{–}10^{-10}$, which suggest that $\epsilon_{\mathrm{a}} = 6.79 \times 10^{-5}$ and $\epsilon_{\mathrm{p}} = 9.74 \times 10^{-5}$. We consider six standard deviations in the statistical fluctuation analysis of the weak + vacuum protocol.

As in the previous subsections, we set $M_{\mathrm{B}} = 10^8$ at the source in Bob's lab. A plug&play setup is assumed. We set $q = 1\%$ to improve the passive estimate efficiency. The imperfections of the intensity monitor are set as follows: the efficiency is set as $\eta_{\mathrm{IM}} = 0.7$ and the noise is set constant as $\sigma_{\mathrm{IM}} = 10^5$. The intensity monitor conservative interval is set constant as $\varsigma = 6\sigma_{\mathrm{IM}} = 6 \times 10^5$.

The simulation results for the weak + vacuum protocol [10] are shown in figure 14. We can see that finite data size clearly reduces the efficiencies of both active and passive estimates. The aforementioned two consequences of finite data size contribute to this efficiency reduction: Firstly, ϵ is nonzero in this finite data size case. Therefore, the estimate of the lower bound of untagged bits' gain is worse as reflected in equation (11). Note that ϵ has the same weight as Δ in equation (11). Secondly, the statistical fluctuation for the weak + vacuum protocol becomes important [14]. Moreover, the tightness of bounds suggested in lemma 1, lemma 2 and proposition 1 may also affect the estimate efficiency in finite data size.

As we showed in section 5.5, using a very bright source can improve the efficiencies of both passive and active estimates. Here we again adjust the source intensity in Bob's lab as $M_{\mathrm{B}} = 10^{10}$. The results are shown in figure 15. We can see that using a very bright source can improve the efficiencies of both passive and active estimates in the finite data size case.

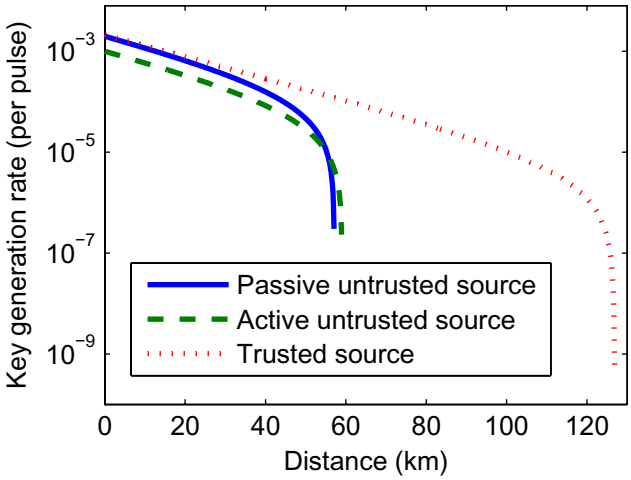

Figure 14. Simulation results of the weak + vacuum [10] protocol with finite data size, asymmetric beam splitter, imperfect intensity monitor and bi-directional structure. We assume that the data size is 10^{12}, the intensity monitor efficiency $\eta_{\mathrm{IM}} = 0.7$, the intensity monitor noise $\sigma_{\mathrm{IM}} = 10^5$, the intensity monitor conservative interval $\varsigma = 6 \times 10^5$, the source in Bob's lab is Poissonian centered at $M_{\mathrm{B}} = 10^8$ photons per pulse and the beam splitting ratio $q = 0.01$. Confidence level is set as $\tau \geqslant 1 - 10^{-10}$. Six standard deviations are considered in the statistical fluctuation. See experimental parameters in table 1. Comparing with figure 11, we can see that finite data size reduces efficiencies of both active and passive estimates.

As we mentioned in section 5.5, such brightness (10^{10} photons per pulse) is achievable with a pulsed laser diode and a fiber laser amplifier. However, nonlinear effects should be carefully considered [30].

In future studies, it would be worthwhile to incorporate the finite key length security analyses [33]–[37] in the key generation rate calculation.

5.7. Simulating the setup in [28]

Peng *et al* [28] report so far the only experimental implementation of QKD that considers the untrusted source imperfection. However, as we discussed above, the analysis proposed in [28] is challenging to use, and was not applied to analyze the experimental results reported in the same paper. Our analysis, however, provides a method to understand the experimental results of [28]. Here, we present a numerical simulation of the system used in [28].

We have to characterize the noise and conservative interval of the intensity monitor used in [28]. The experimental results reported in [28] show that the *measured* input photon number distribution is centered at $M = 1.818 \times 10^7$ with a standard deviation 3.097×10^5 on Alice's side. If we assume the source at Bob's side to be Poissonian, the *actual* input photon number distribution on Alice's side will also be Poissonian. The detector noise is then $\sigma_{\mathrm{IM}} = \sqrt{(3.097 \times 10^5)^2 - 1.818 \times 10^7} = 3.097 \times 10^5$. We set the detector conservative interval as constant $\varsigma = 6\sigma_{\mathrm{IM}}$.

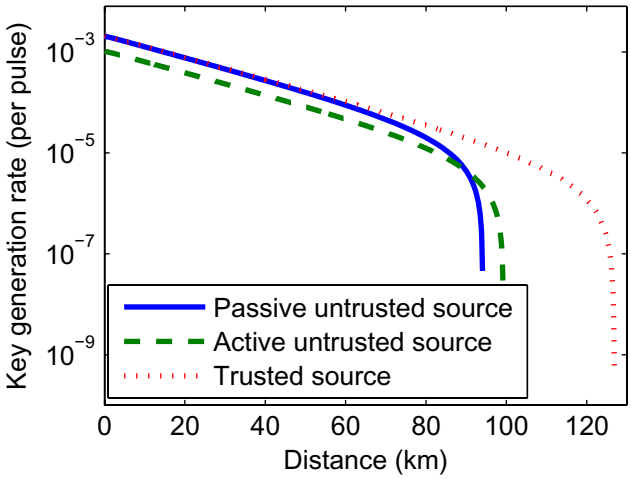

Figure 15. Simulation result of the weak + vacuum [10] protocol with finite data size, asymmetric beam splitter, imperfect intensity monitor, bi-directional structure and very bright source. We assume that the data size is 10^{12}, the intensity monitor efficiency $\eta_{\mathrm{IM}} = 0.7$, the intensity monitor noise $\sigma_{\mathrm{IM}} = 10^5$, the intensity monitor conservative interval $\varsigma = 6 \times 10^5$, the source in Bob's lab is Poissonian centered at $M_{\mathrm{B}} = 10^{10}$ photons per pulse and the beam splitting ratio $q = 0.01$. Confidence level is set as $\tau \geqslant 1 - 10^{-10}$. Six standard deviations are considered in the statistical fluctuation. Citing experimental parameters from table 1. Comparing with figure 11, we can see that using a very bright source can improve efficiencies of both active and passive estimates.

Source intensity at Bob's side M_{B} can be calculated in the following manner: since $M = 1.818 \times 10^7$ at a distance $l = 25$ km and beam splitting ratio $q = 0.05$, we can conclude that

$$M_{\mathrm{B}} = \frac{M}{\alpha l (1-q)} = 6.411 \times 10^7.$$

Here we assume that the fiber loss coefficient $\alpha = -0.21\ \mathrm{dB\ km^{-1}}$.

The other parameters are directly cited from [28]: the setup is in plug&play structure. The efficiency of the intensity monitor is $\eta_{\mathrm{IM}} = 0.8$. Single photon detector efficiency is 4%, detector error rate is 1.39% and background rate $Y_0 = 9.38 \times 10^{-5}$. As in the previous sections, confidence level is set as $\tau \geqslant 1\text{–}10^{-10}$.

In the experiment reported in [28], the data size is 9.05×10^7. (It is smaller than the data size we assumed in other simulations. If a larger data size were used, we would expect some improvements in the simulation results.) We ran numerical simulation with six standard deviations that are considered in the statistical fluctuation. The simulation results are shown in figure 16. It is encouraging to see that the simulation yields positive key rates for both passive and active estimates at short distances.

5.8. Summary

From the numerical simulations shown in figures 5–16, we conclude that four important parameters can improve the efficiency of passive estimate on an untrusted source: Firstly, the

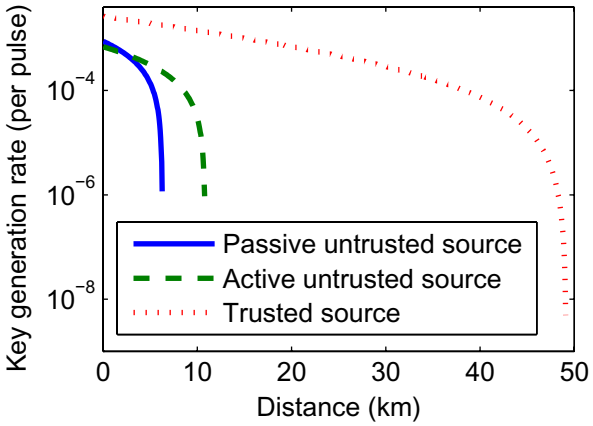

Figure 16. Simulation result of the weak + vacuum [10] protocol based on the experimental parameters in [28]: data size is 9.05×10^7 (this data size is reported in [28]. It is smaller than the data size we assumed in other simulations. If a larger data size was used, we would expect some improvements on the simulation results), the intensity monitor efficiency $\eta_{\mathrm{IM}} = 0.8$, the intensity monitor noise $\sigma_{\mathrm{IM}} = 3.097 \times 10^5$, the intensity monitor conservative interval $\varsigma = 6\sigma_{\mathrm{IM}}$, the source at Bob's side is Poissonian centered at $M_{\mathrm{B}} = 6.411 \times 10^7$ photons per pulse, the beam splitting ratio $q = 0.05$ and the system is in plug&play. Confidence level is set as $\tau \geqslant 1\text{–}10^{-10}$. Six standard deviations are considered in the statistical fluctuation. Single photon detector efficiency is 4%, detector error rate is 1.39% and background rate $Y_0 = 9.38 \times 10^{-5}$. Comparing with figure 14, we can see that a higher background rate limits the system performance.

beam splitting ratio q should be very small, say 1%, to send most input photons to the intensity monitor. Secondly, the light source should be very bright (say, 10^{10} photons per pulse). This is particularly important for plug&play structure. Thirdly, the imperfections of the intensity monitor should be small. That is, the intensity monitor should have high efficiency (say, over 70%) and high precision (say, can resolve photon number difference of 6×10^5). Fourthly, the data size should be large (say, 10^{12} bits) to minimize the statistical fluctuation.

In brief, a largely biased beam splitter, a bright source, an efficient and precise intensity monitor and a large data size are four key conditions that can substantially improve the efficiency of the passive estimate on an untrusted source. The latter three conditions are also applicable to the active estimate scheme.

An important advantage of decoy state protocols is that the key generation rate will only drop linearly as channel transmittance decreases [7]–[14], whereas in many non-decoy protocols, like the GLLP protocol [5], the key generation rate will drop quadratically as channel transmittance decreases. In the simulations shown in figures 6–16, we can see that this important advantage is preserved even if the source is unknown and untrusted.

6. Preliminary experimental test

We performed some preliminary experiments to test our analysis. The basic idea is to measure some key parameters of our system, especially the characteristics of the source, with which we can perform numerical simulation to show the expected performance.

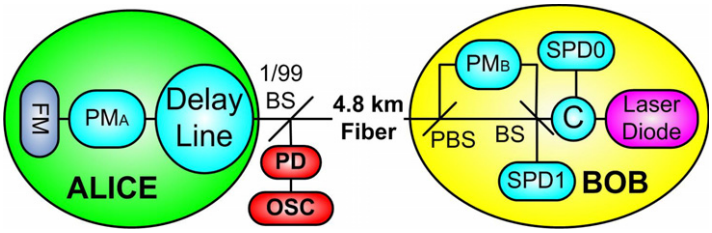

Figure 17. Experimental setup. Alice and Bob: commercial plug&play QKD system. PD: photodiode. OSC: high-speed oscilloscope. 1/99 BS: 1/99 beam splitter. FM: Faraday mirror. PM_x: phase modulators. PBS: polarizing beam splitter. BS: beam splitter. SPD_x: single photon detector. C: circulator.

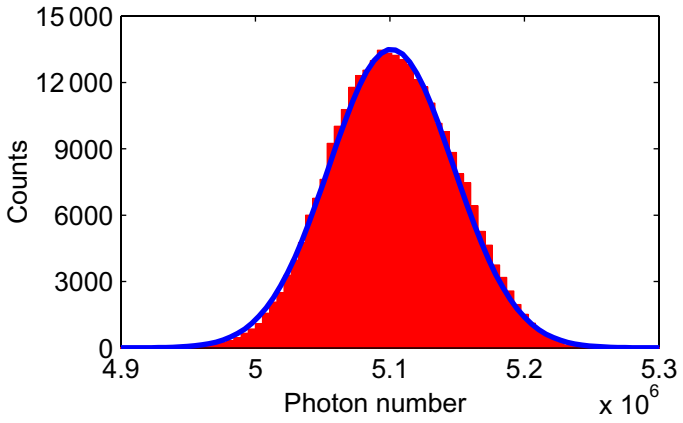

Figure 18. Experimentally measured photon number statistics for 299 700 pulses. The distribution centered at 5.101×10^6 photons per pulse, with standard deviation 6.557×10^4. The blue line shows a Gaussian fit of the actual distribution.

The experimental setup is shown in figure 17. It is essentially a modified commercial plug&play QKD system. We added a 1/99 beam splitter (1/99 BS in figure 17), a photodiode (PD in figure 17) and a high-speed oscilloscope (OSC in figure 17) on Alice's side. These three parts constitute Alice's PNA.

When Bob sends strong laser pulses to Alice, the photodiode (PD in figure 17) will convert input photons into photoelectrons, which are then recorded by the oscilloscope (OSC in figure 17). In the recorded waveform, we calculated the area below each pulse. This area is proportional to the number of input photons. The conversion coefficient between the area and photon number is calibrated by measuring the average input laser power on Alice's side with a slow optical power meter.

In our experiment, 299 700 pulses are generated by the laser diode at Bob's side (laser diode in figure 17) at a repetition rate of 5 MHz with 1 ns pulse width. They are all split into U pulses and L pulses (see figure 4) by the 1/99 beam splitter (1/99 BS in figure 17). The L pulses are measured by a photodiode (PD in figure 17). The measurement results are acquired and recorded by an oscilloscope (OSC in figure 17).

The experimental results of the photon number statistics are plotted in figure 18. The measured photon number distribution is centered at $M = 5.101 \times 10^6$ photons per pulse, with

Table 2. Parameters measured from our preliminary experiment described in section 6.

α	η_{det}	e_{det}	Y_0
$-0.21\ \text{dB km}^{-1}$	4.89%	0.21%	8.4×10^{-5}

standard deviation 6.557×10^4 on Alice's side. We can see that the actual photon number distribution fits a Gaussian distribution (shown as the blue line) well. Other experimental results are shown in table 2.

The intensity monitor noise is calculated in a similar manner to that in section 5.7: assuming the source is Poissonian at Bob's side, which means the actual input photon number on Alice's side is also Poissonian, the noise is then given by $\sigma_{\text{IM}} = \sqrt{(6.557 \times 10^4)^2 - 5.101 \times 10^6} = 6.553 \times 10^4$. As in section 5.7, we set the detector conservative interval as a constant $\varsigma = 6\sigma_{\text{IM}}$.

Source intensity at Bob's side M_{B} can be calculated in the following matter (which is similar to the one we used in section 5.7): since $M = 5.101 \times 10^6$ at a distance $l = 4.8$ km, and beam splitting ratio $q = 0.01$, we can conclude that

$$M_{\text{B}} = \frac{M}{\alpha l(1-q)} = 6.500 \times 10^6.$$

Here we know that the fiber loss coefficient $\alpha = -0.21\ \text{dB km}^{-1}$.

The simulation result is shown in figure 19, in which the data size is set as 10^{12}.[8] We can see that it is possible to achieve positive key rate at moderate distances using the security analysis presented in this paper.

7. Conclusion

In this paper, we present the first passive security analysis for QKD with an untrusted source, with a complete security proof. Our proposal is compatible with inefficient and noisy intensity monitors, which is not considered in [21] or in [28]. Our analysis is also compatible with a finite data size, which is not considered in [28]. Comparing with the active estimate scheme proposed in [21], the passive scheme proposed in this paper significantly reduces the challenges to implement the 'plug&play' QKD with unconditional security. Our proposal can be applied to practical QKD setups with untrusted sources, especially the plug&play QKD setups, to guarantee the security.

We point out four important conditions that can improve the efficiency of the passive estimate scheme proposed in this paper: firstly, the beam splitter in PNA should be largely biased to send most photons to the intensity monitor. Secondly, the light source should be bright. Thirdly, the intensity monitor should have high efficiency and precision. Fourthly, the data size should be large to minimize the statistical fluctuation. These four conditions are confirmed in extensive numerical simulations.

[8] Data size in our experiment is much smaller than the data size assumed in numerical simulation. The purpose of our preliminary experiment is to test whether it is possible to achieve positive key rate with our current system.

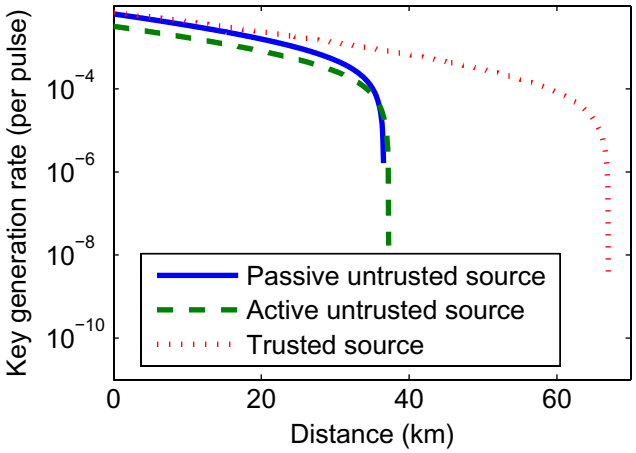

Figure 19. Simulation result of the weak + vacuum [10] protocol based on experimental parameters from our QKD system. We assume that the data size is 10^{12} bits, the intensity monitor efficiency $\eta_{\mathrm{IM}} = 0.7$, the intensity monitor noise $\sigma_{\mathrm{IM}} = 6.553 \times 10^4$, the intensity monitor conservative interval $\varsigma = 6\sigma_{\mathrm{IM}}$, the source at Bob's lab is Poissonian centered at $M_{\mathrm{B}} = 6.500 \times 10^6$ photons per pulse, the beam splitting ratio $q = 0.01$ and the system is in the plug&play structure. Confidence level is set as $\tau \geqslant 1 - 10^{-10}$. Six standard deviations are considered in the statistical fluctuation. Experimental parameters are listed in table 2.

In the simulations shown in figures 11–16 and 19, we made an additional assumption that the intensity monitor has a constant Gaussian noise. This assumption is *not* required by our security analysis. It will be interesting to experimentally verify this model in future.

The numerical simulations show that if the above conditions are met, the efficiency of the passive untrusted source estimate is close to that of the trusted source estimate, and is roughly twice as high as the efficiency of the active untrusted source estimate. Nonetheless, the efficiency of the active estimate scheme proposed in [21] may be improved to the level that it is similar to the efficiency of passive estimation. This is briefly discussed below in equation (C.3). The security of the improved active estimate scheme is beyond the scope of the current paper, and is a subject for further investigation.

Numerical simulations in figures 6–16 and 19 show that the key generation rate drops linearly as the channel transmittance decreases. This is an important advantage of decoy state protocols over many other QKD protocols, and is preserved in our untrusted source analysis.

Our preliminary experimental test highlights the feasibility of our proposed passive estimate scheme. Indeed, our scheme can be easily implemented by making very simple modifications (by adding a few commercial modules) to a commercial plug&play QKD system.

A remaining practical question in our proposal is: How to calibrate the noise and the conservative interval of the intensity monitor? Note that these two parameters may not be constant at different intensity levels. Moreover, the noise may not be Gaussian. It is not straightforward to define the conservative interval and its confidence.

Acknowledgments

We thank Lei Wu and Marc Napoleon for performing some experimental tests, which were supported by NSERC USRA. We also thank Jean-Christian Boileau for discussions in the early stage of this work. Support from the funding agencies CFI, CIPI, the CRC program, CIFAR, MITACS, NSERC, OIT, PREA and Quantum Works is gratefully acknowledged.

Appendix A. The phase randomization assumption

In this section, we will show that for the state that is accessible to Eve, Alice's phase randomization is equivalent to performing a QND measurement on the photon number of the input pulses.

Proof. Before the phase randomization, the state that is shared by Alice, Bob and Eve is

$$|\psi\rangle_{\mathrm{ABE}} = \sum_{m,n} b_{m,n}|E_n\rangle|m\rangle. \tag{A.1}$$

After the phase randomization, a random phase is applied to each pulse, and this phase is inaccessible to Eve. The state becomes

$$|\psi\rangle^{\mathrm{pr}}_{\mathrm{ABE}} = \sum_{m,n} b_{m,n}|E_n\rangle|m\rangle \mathrm{e}^{\mathrm{i}m\theta}. \tag{A.2}$$

Its density matrix is

$$\rho^{\mathrm{pr}}_{\mathrm{ABE}} = \sum_{m,n,m',n'} b_{m,n}b^*_{m',n'}|E_n\rangle\langle E_{n'}| \otimes |m\rangle\langle m'|\mathrm{e}^{\mathrm{i}(m-m')\theta}. \tag{A.3}$$

Since θ is not known to Eve or Bob, the state that is accessible to Bob and Eve is

$$\begin{aligned}\rho^{\mathrm{pr}}_{\mathrm{BE}} &= \frac{1}{2\pi}\sum_{m,n,m',n'} b_{m,n}b^*_{m',n'}|E_n\rangle\langle E_{n'}| \otimes |m\rangle\langle m'| \int_0^{2\pi} \mathrm{e}^{\mathrm{i}(m-m')\theta}\mathrm{d}\theta \\ &= \sum_{m,n,n'} b_{m,n}b^*_{m,n'}|E_n\rangle\langle E_{n'}| \otimes |m\rangle\langle m|. \end{aligned} \tag{A.4}$$

Instead of considering the phase randomization, now let us analyze the impact of a photon number measurement on the state given in equation (A.1). Before Alice's measurement, the density matrix is given by

$$\rho_{\mathrm{ABE}} = \sum_{m,n,m',n'} b_{m,n}b^*_{m',n'}|E_n\rangle\langle E_{n'}| \otimes |m\rangle\langle m'|. \tag{A.5}$$

After the QND measurement of the photon number, Alice knows the photon number. However, Eve and Bob do not know Alice's measurement result. Therefore, the state that is accessible to Bob and Eve is given by

$$\begin{aligned}\rho^{\mathrm{QND}}_{\mathrm{BE}} &= \sum_{m\prime} |m'\rangle\langle m'|\rho_{\mathrm{ABE}}|m'\rangle\langle m'| \\ &= \sum_{m,n,n'} b_{m,n}b^*_{m,n'}|E_n\rangle\langle E_{n'}| \otimes |m\rangle\langle m| \\ &= \rho^{\mathrm{pr}}_{\mathrm{BE}}. \end{aligned} \tag{A.6}$$

□

From the above equation, we conclude that the two different processes: (i) phase randomization by Alice and (ii) a photon-number non-demolition measurement, actually give exactly the same density matrix for the Eve–Bob system. Therefore, phase randomization is mathematically equivalent to a photon-number non-demolition measurement. For this reason, we can consider the output state by Alice as a classical mixture of Fock states.

Appendix B. Security analysis for untagged bits

In this section, for the convenience of the readers, we recapitulate the security analysis that is presented in [21].

Assume that k pulses are sent from Alice to Bob. Alice and Bob do not know which bits are untagged. However, either from the active estimate presented in [21] or from the passive estimate presented in the current paper, they know that at least $(1-\Delta-\epsilon)k$ pulses are untagged with high confidence.

Alice and Bob can measure the overall gain $Q_{\rm e}$ and the overall QBER $E_{\rm e}$. They do not know the gain Q and the QBER E for the untagged bits because they do not know which bits are untagged. Nonetheless, they can then estimate the upper bounds and the lower bounds of them. The upper bound and the lower bound of Q are [21]

$$\begin{aligned} \overline{Q} &= \frac{Q_{\rm e}}{1-\Delta-\epsilon}, \\ \underline{Q} &= \max\left(0, \frac{Q_{\rm e}-\Delta-\epsilon}{1-\Delta-\epsilon}\right). \end{aligned} \tag{B.1}$$

The upper bound and the lower bound of $E\cdot Q$ can be estimated as [21]

$$\begin{aligned} \overline{E\cdot Q} &= \frac{Q_{\rm e}E_{\rm e}}{1-\Delta-\epsilon}, \\ \underline{E\cdot Q} &= \max\left(0, \frac{Q_{\rm e}E_{\rm e}-\Delta-\epsilon}{1-\Delta-\epsilon}\right). \end{aligned} \tag{B.2}$$

For untagged bits (i.e. $m\in[(1-\delta)M,(1+\delta)M]$), we can show that the upper bound and the lower bound of the probability that the output photon number from Alice is n are [21]

$$\begin{aligned} \overline{P_n} &= \begin{cases} (1-\lambda)^{(1-\delta)M} & \text{if } n=0, \\ \binom{(1+\delta)M}{n}\lambda^n(1-\lambda)^{(1+\delta)M-n} & \text{if } 1\leqslant n\leqslant(1+\delta)M, \\ 0 & \text{if } n>(1+\delta)M, \end{cases} \\ \underline{P_n} &= \begin{cases} (1-\lambda)^{(1+\delta)M} & \text{if } n=0, \\ \binom{(1-\delta)M}{n}\lambda^n(1-\lambda)^{(1-\delta)M-n} & \text{if } 1\leqslant n\leqslant(1-\delta)M, \\ 0 & \text{if } n>(1-\delta)M, \end{cases} \end{aligned} \tag{B.3}$$

under **Condition 1**:

$$(1+\delta)M\lambda < 1. \tag{B.4}$$

The key rate calculation depends on the QKD protocol that is implemented. For the GLLP [5] protocol with an untrusted source, the key generation rate is given by [21]

$$R \geqslant \frac{1}{2}\left\{-Q_{\mathrm{e}} f(E_{\mathrm{e}}) H_2(E_{\mathrm{e}}) + (\underline{Q} + \underline{P_0} + \overline{P_1} - 1)\left[1 - H_2\left(\frac{Q_{\mathrm{e}} E_{\mathrm{e}}}{\underline{Q} + \underline{P_0} + \overline{P_1} - 1}\right)\right]\right\}, \tag{B.5}$$

where Q_{e} and E_{e} are measured experimentally, $\underline{Q}$ can be calculated from equation (B.1), and $\underline{P_0}$ and $\overline{P_1}$ can be calculated from equation (B.3).

For decoy state protocols [7]–[14], the key generation rate (with an untrusted source) is given by [21]

$$R \geqslant \tfrac{1}{2}\{-Q_{\mathrm{e}}^{\mathrm{S}} f(E_{\mathrm{e}}^{\mathrm{S}}) H_2(E_{\mathrm{e}}^{\mathrm{S}}) + (1 - \Delta - \epsilon)\underline{Q_1^{\mathrm{S}}}[1 - H_2(\overline{e_1^{S}})]\}, \tag{B.6}$$

where $Q_{\mathrm{e}}^{\mathrm{S}}$ and $E_{\mathrm{e}}^{\mathrm{S}}$ are the overall gain and the over QBER of the signal states, respectively, and can be measured experimentally. $\underline{Q_1^{\mathrm{S}}}$ and $\overline{e_1^{\mathrm{S}}}$ depend on the specific decoy state protocol that is implemented.

For the weak + vacuum protocol [10, 11, 14], the lower bound of Q_1^{S} for untagged bits is given by [21]

$$\underline{Q_1^{\mathrm{S}}} = \underline{P_1^{\mathrm{S}}}\frac{\underline{Q^{\mathrm{D}}}\,\underline{P_2^{\mathrm{S}}} - \overline{Q^{\mathrm{S}}}\,\overline{P_2^{\mathrm{D}}} + (\underline{P_0^{\mathrm{S}}}\,\overline{P_2^{\mathrm{D}}} - \overline{P_0^{\mathrm{D}}}\,\underline{P_2^{\mathrm{S}}})\overline{Q^{\mathrm{V}}} - \frac{2\delta M(1-\lambda_{\mathrm{D}})^{2\delta M-1}\underline{P_2^{\mathrm{S}}}}{[(1-\delta)M+1]!}}{\overline{P_1^{\mathrm{D}}}\,\underline{P_2^{\mathrm{S}}} - \underline{P_1^{\mathrm{S}}}\,\overline{P_2^{\mathrm{D}}}} \tag{B.7}$$

under **Condition 2:**

$$\frac{\lambda_{\mathrm{S}}}{\lambda_{\mathrm{D}}} > \frac{(1+\delta)M-2}{(1-\delta)M-2}\left[\frac{(1+\delta)M-2}{2\delta M}\right]^{2\delta M/((1-\delta)M-2)}\left[\frac{(1+\delta)M-2}{(1-\delta)M-2}\cdot\frac{e^2}{2\delta M}\right]^{1/(2[(1-\delta)M-2])}. \tag{B.8}$$

Here Q^{S}, Q^{D} and Q^{V} are the gains of untagged bits of the signal state, the decoy state and the vacuum state, respectively. Their bounds can be estimated from equations (B.1). The bounds of the probabilities can be estimated from equations (B.3). λ_{S} and λ_{D} are Alice's internal transmittances for signal and decoy states, respectively. The upper bound of e_1^{S} for untagged bits is given by [21]

$$e_1^{\mathrm{S}} \leqslant \overline{e_1^{\mathrm{S}}} = \frac{\overline{E^{\mathrm{S}}Q^{\mathrm{S}}} - \underline{P_0^{\mathrm{S}}}\,\underline{E^{\mathrm{V}}Q^{\mathrm{V}}}}{\underline{Q_1^{\mathrm{S}}}}, \tag{B.9}$$

in which E^{S} and E^{V} are the QBERs of untagged bits of the signal and the vacuum states, respectively. $\overline{E^{\mathrm{S}}Q^{\mathrm{S}}}$ and $\underline{E^{\mathrm{V}}Q^{\mathrm{V}}}$ can be estimated from equations (B.2). $\underline{P_0^{\mathrm{S}}}$ can be estimated by equations (B.3). $\underline{Q_1^{\mathrm{S}}}$ is given by equation (B.7). $\underline{E^{\mathrm{V}}Q^{\mathrm{V}}}$ can be estimated from equations (B.2). $\underline{P_0^{\mathrm{S}}}$ can be estimated by equations (B.3). $\underline{Q_1^{\mathrm{S}}}$ is given by equation (B.7).

For one-decoy protocol [10, 13], a lower bound of Q_1^S and an upper bound of e_1^S for untagged bits are given by

$$\underline{Q_1^S} = \underline{P_1^S}\frac{\underline{Q^D P_2^S} - \overline{Q^S}\,\overline{P_2^D} + (\underline{P_0^S}\,\overline{P_2^D} - \overline{P_0^D}\,\underline{P_2^S})\frac{\overline{E^S Q^S}}{\underline{P_0^S} E^V} - \frac{2\delta M(1-\lambda_D)^{2\delta M-1}\underline{P_2^S}}{[(1-\delta)M+1]!}}{\overline{P_1^D}\,\underline{P_2^S} - \underline{P_1^S}\,\overline{P_2^D}},$$
$$\overline{e_1^S} = \frac{\overline{E^S \cdot Q^S}}{\underline{Q_1^S}}, \tag{B.10}$$

respectively, under condition 2 in the asymptotic case. Here Q^S and Q^D are the gains of untagged bits of the signal state and the decoy state, respectively. Their bounds can be estimated from equations (B.1). E^S is the QBER of untagged bits of the signal state. $\overline{E^S \cdot Q^S}$ can be estimated from equations (B.2). $E^V = 0.5$ in the asymptotic case. The bounds of the probabilities can be estimated from equations (B.3).

Appendix C. Confidence level in active estimate

Among all the V untagged bits, each bit has probability 1/2 to be assigned as an untagged coding bit. Therefore, the probability that $V_c = v_c$ obeys a binomial distribution. Cumulative probability is given by [38]

$$P\left(V_c \leqslant \frac{V-\epsilon k}{2} | V = v\right) \leqslant \exp\left(-\frac{\epsilon^2 k^2}{2v}\right).$$

For any $v \in [0, k]$, $k/v \geqslant 1$. Therefore, we have

$$P\left(V_c \leqslant \frac{V-\epsilon k}{2} | V \in [0, k]\right) \leqslant \exp\left(-\frac{k\epsilon^2}{2}\right).$$

In the experiment described by lemma 1, $V \in [0, k]$ is always true. Therefore, the above inequality reduces to

$$P\left(V_c \leqslant \frac{V-\epsilon k}{2}\right) \leqslant \exp\left(-\frac{k\epsilon^2}{2}\right). \tag{C.1}$$

By definition, we have

$$V = V_c + V_s. \tag{C.2}$$

Substituting equation (C.2) into equation (C.1), we have

$$P(V_c \leqslant V_s - \epsilon k) \leqslant \exp\left(-\frac{k\epsilon^2}{2}\right). \tag{C.3}$$

The above proof can be easily generalized to the case where for each bit sent from the untrusted source to Alice, Alice randomly assigns it as either a coding bit with probability γ, or a sampling bit with probability $1-\gamma$. Here $\gamma \in (0, 1)$ is chosen by Alice. It is then straightforward to show that

$$P\left[V_c \leqslant \frac{\gamma}{1-\gamma}(V_s - \epsilon k)\right] \leqslant \exp(-2k\epsilon^2\gamma^2). \tag{C.4}$$

When $\gamma = 1/2$, equation (C.4) reduces to equation (C.3).

Appendix D. Confidence level in cross estimate

From corollaries 3 and 5, we know that

$$P(V_{\mathrm{c}}^{\mathrm{U}} \leqslant V_{\mathrm{s}}^{\mathrm{L}} - \epsilon_1 k) \leqslant \exp\left(\frac{-k\epsilon_1^2}{2}\right),$$
$$P(V_{\mathrm{s}}^{\mathrm{U}} \leqslant V_{\mathrm{c}}^{\mathrm{L}} - \epsilon_2 k) \leqslant \exp\left(\frac{-k\epsilon_2^2}{2}\right). \tag{D.1}$$

Therefore, we have

$$\begin{aligned} P(V^{\mathrm{U}} \leqslant V_{\mathrm{s}}^{\mathrm{L}} + V_{\mathrm{c}}^{\mathrm{L}} - (\epsilon_1+\epsilon_2)k) &= P(V_{\mathrm{c}}^{\mathrm{U}} + V_{\mathrm{s}}^{\mathrm{U}} \leqslant V_{\mathrm{s}}^{\mathrm{L}} + V_{\mathrm{c}}^{\mathrm{L}} - (\epsilon_1+\epsilon_2)k) \\ &\leqslant P[(V_{\mathrm{c}}^{\mathrm{U}} \leqslant V_{\mathrm{s}}^{\mathrm{L}} - \epsilon_1 k) \text{ or } (V_{\mathrm{s}}^{\mathrm{U}} \leqslant V_{\mathrm{c}}^{\mathrm{L}} - \epsilon_2 k)] \\ &\leqslant P(V_{\mathrm{c}}^{\mathrm{U}} \leqslant V_{\mathrm{s}}^{\mathrm{L}} - \epsilon_1 k) + P(V_{\mathrm{c}}^{\mathrm{U}} \leqslant V_{\mathrm{s}}^{\mathrm{L}} - \epsilon_2 k) \\ &= \exp\left(\frac{-k\epsilon_1^2}{2}\right) + \exp\left(\frac{-k\epsilon_2^2}{2}\right). \end{aligned} \tag{D.2}$$

In the above derivation, we made use of the fact that $[(V_{\mathrm{c}}^{\mathrm{U}} \leqslant V_{\mathrm{s}}^{\mathrm{L}} - \epsilon_1 k) \text{ or } (V_{\mathrm{s}}^{\mathrm{U}} \leqslant V_{\mathrm{c}}^{\mathrm{L}} - \epsilon_2 k)]$ is always true if $[V_{\mathrm{c}}^{\mathrm{U}} + V_{\mathrm{s}}^{\mathrm{U}} \leqslant V_{\mathrm{s}}^{\mathrm{L}} + V_{\mathrm{c}}^{\mathrm{L}} - (\epsilon_1+\epsilon_2)k]$ is true.

References

[1] Bennett C H and Brassard G 1984 Quantum cryptography: public key distribution and coin tossing *Proc. IEEE Int. Conf. on Computers, Systems and Signal Processing* pp 175–9

[2] Ekert A K 1991 Quantum cryptography based on Bell's theorem *Phys. Rev. Lett.* **67** 661

[3] Lo H-K and Zhao Y 2009 Quantum cryptography *Encyclopedia of Complexity and System Science* vol 8 (New York: Springer) pp 7265–89 (arXiv:0803.2507)

[4] Mayers D 2001 *J. ACM* **48** 351
Lo H-K and Chau H F 1999 *Science* **283** 2050
Shor P and Preskill J 2000 *Phys. Rev. Lett.* **85** 441

[5] Gottesman D, Lo H-K, Lütkenhaus N and Preskill J 2004 Security of quantum key distribution with imperfect devices *Quantum. Inf. Comput.* **4** 325

[6] Inamori H, Lütkenhaus N and Mayers D 2007 Unconditional security of practical quantum key distribution *Eur. Phys. J.* D **41** 599

[7] Hwang W Y 2003 Quantum key distribution with high loss: toward global secure communication *Phys. Rev. Lett.* **91** 057901

[8] Lo H-K 2004 Quantum key distribution with vacua or dim pulses as decoy states *Proc. IEEE Int. Symp. on Information Theory* p 137

[9] Lo H-K, Ma X and Chen K 2005 Decoy state quantum key distribution *Phys. Rev. Lett.* **94** 230504

[10] Ma X, Qi B, Zhao Y and Lo H-K 2005 Practical decoy state for quantum key distribution *Phys. Rev.* A **72** 012326

[11] Wang X-B 2005 Beating the photon-number-splitting attack in practical quantum cryptography *Phys. Rev. Lett.* **94** 230503

[12] Wang X-B 2005 Decoy-state protocol for quantum cryptography with four different intensities of coherent light *Phys. Rev.* A **72** 012322

[13] Zhao Y, Qi B, Ma X, Lo H-K and Qian L 2006 Experimental quantum key distribution with decoy states *Phys. Rev. Lett.* **96** 070502

[14] Zhao Y, Qi B, Ma X, Lo H-K and Qian L 2006 Simulation and implementation of decoy state quantum key distribution over 60 km telecom fiber *Proc. IEEE Int. Symp. on Information Theory* pp 2094–8

[15] Muller A, Herzog T, Hutter B, Tittel W, Zbinden H and Gisin N 1997 'Plug&play' systems for quantum crytography *Appl. Phys. Lett.* **70** 793
[16] Stucki D, Gisin N, Guinnard O, Ribordy G and Zbinden H 2002 Quantum key distribution over 67 km with a plug&play system *New J. Phys.* **4** 41
[17] www.idquantique.com
[18] www.magiqtech.com
[19] Pólya G 1920 Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem (in German) *Math. Z.* **8** 171
[20] Gisin N, Fasel S, Kraus B, Zbinden H and Ribordy G 2006 Trojan horse attacks on quantum key distribution systems *Phys. Rev.* A **73** 022320
[21] Zhao Y, Qi B and Lo H-K 2008 Quantum key distribution with an unknown and untrusted source *Phys. Rev.* A **77** 052327
[22] Villoresi P *et al* 2008 Experimental verification of the feasibility of a quantum channel between space and earth *New J. Phys.* **10** 033038
[23] Wang X-B, Peng C-Z and Pan J-W 2007 Simple protocol for secure decoy-state quantum key distribution with a loosely controlled source *Appl. Phys. Lett.* **90** 031110
[24] Wang X-B, Peng C-Z, Zhang J and Pan J-W 2008 Security of decoy-state quantum key distribution with inexactly controlled source arXiv:quant-ph/0612121v3
[25] Wang X-B, Peng C-Z, Zhang J and Pan J-W 2008 General theory of decoy-state quantum cryptography with source errors *Phys. Rev.* A **77** 042311
[26] Wang X-B, Yang L, Peng C-Z and Pan J-W 2009 Decoy-state quantum key distribution with both source errors and statistical fluctuations *New J. Phys.* **11** 075006
[27] Takesue H, Nam S W, Zhang Q, Hadfield R H, Honjo T, Tamaki K and Yamamoto Y 2007 Quantum key distribution over a 40-dB channel loss using superconducting single-photon detectors *Nat. Photonics* **1** 343
[28] Peng X, Jiang H, Xu B, Ma X and Guo H 2008 Experimental quantum key distribution with an untrusted source *Opt. Lett.* **33** 2077
[29] Gobby C, Yuan Z L and Shields A J 2004 Quantum key distribution over 122 km of standard telecom fiber *Appl. Phys. Lett.* **84** 3762
[30] See, e.g., Keiser G 2000 *Optical Fiber Communications* 3rd edn (New York: McGraw-Hill) chapter 12.5
[31] Zhang Q *et al* 2009 Megabits secure key rate quantum key distribution *New J. Phys.* **11** 045010
[32] Dixon A R, Yuan Z L, Dynes J F, Sharpe A W and Shields A J 2008 Gigahertz decoy quantum key distribution with 1 Mbit/s secure key rate *Opt. Express* **16** 18790
[33] Hayashi M 2006 Practical evaluation of security for quantum key distribution *Phys. Rev.* A **74** 022306
[34] Hayashi M 2007 Upper bounds of eavesdropper's performances in finite-length code with decoy method *Phys. Rev.* A **76** 012329
[35] Hasegawa J, Hayashi M, Hiroshima T and Tomita A 2007 Security analysis of decoy state quantum key distribution incorporating finite statistics arXiv:0707.3541
[36] Scarani V and Renner R 2008 Quantum cryptography with finite resources: unconditional security bound for discrete-variable protocols with one-way postprocessing *Phys. Rev. Lett.* **100** 200501
[37] Ma X, Fung C-H F, Boileau J-C and Chau H F 2009 Practical post-processing for quantum-key-distribution experiments arXiv:0904.1994
[38] Hoeffding W 1963 Probability inequalities for sums of bounded random variables *J. Am. Stat. Assoc.* **58** 13